Technical Attachment

Comparative Verification for Dallas Fort Worth and Waco

Mark A. Fox and Steven J. Fano
WFO Forth Worth, Texas

1. Introduction

There are a growing number of sources where people can obtain weather forecasts. These outlets include, but are not limited to, the National Weather Service, local and cable television stations, private weather companies, as well as model derived forecast (MOS forecasts). The purpose of this study is to compare temperature forecasts made by a number of these sources in an effort to rate their accuracy. Only temperatures are compared in this study since there is no uniformity in the use of probability of precipitation forecasts among the private weather companies.

This study was conducted over a 12 month period and examined the temperature forecasts for both the Dallas/Fort Worth Airport (**DFW**) and the Waco Regional Airport (**Waco**). Two forecasts per day were obtained from two private weather companies, two numerical models, and from the National Weather Service office in Fort Worth (**FWD**). Since it was not the intent of this study to make any participant look good or bad, the private weather company names have not been disclosed and will be referenced as Company 1 **(C 1**) and Company 2 **(C 2**) for the remainder of this study.

2. Methodology

The study began on July 26, 2000 and ran continuously until July 26, 2001. Two temperature forecasts per day were verified which corresponded to the 00 UTC AND 12 UTC model runs. Each forecast consisted of four time periods, and included a 12, 24, 36, and 48 hour forecast. Temperature forecasts that were issued by the National Weather Service in Fort Worth were obtained from the coded city forecast (**CCF**) product, issued around 3:30 AM and 3:30 PM each day. The private company's forecasts were acquired from their own web pages where a zip code was entered to obtain a local forecast. The zip codes used were 75261 and 76708, which corresponded to DFW airport and the Waco Regional airport respectively. It should be noted that the FWD forecasts were the first to be released and the private company's forecasts were usually available two hours later. The model output statistics from the Nested Grid Model were obtained from the **FWC** product and the Aviation model output statistics from the **FAN** product. The new MAV guidance from the Aviation model was not available at the time of the study.

Once the actual temperatures were obtained, calculations were made to determine the absolute error for each forecast. This was done by simply subtracting the observed temperature at each station from the forecasted temperature and taking the absolute value of the error. The forecasts and results were then entered into a spread sheet in order to run a statistical analysis. All of the forecasts were also compared to climatological normals (**CLI**) for each calendar day.

3. Results

Tabulated scores for total forecast error for all periods (Fig. 1) show that FWD had the fewest number of forecast errors while climatology had the most. Both private companies had fewer total errors than the FAN guidance, although C1 had more total error than the FWC guidance.

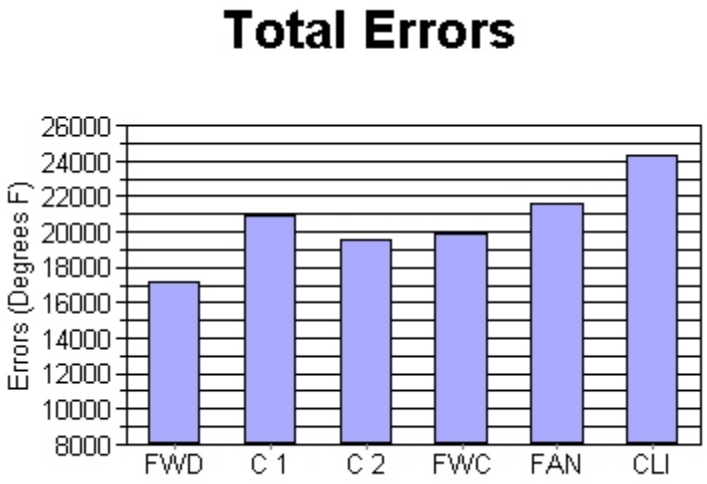


Fig 1. Total errors in degrees Fahrenheit for 732 forecasts. Each forecast contains 4 periods.

When the total errors are divided up into time periods (Figs. 2a and 2b) it becomes apparent that the majority of the errors occurred in period four (P4) for all participants. FWD had the fewest errors in each period at both DFW and Waco. The FWC scored better than C1 in all periods at DFW and in period one and three at Waco. C2 had fewer errors than the FWC in periods two and four at DFW and in period three and four at Waco.

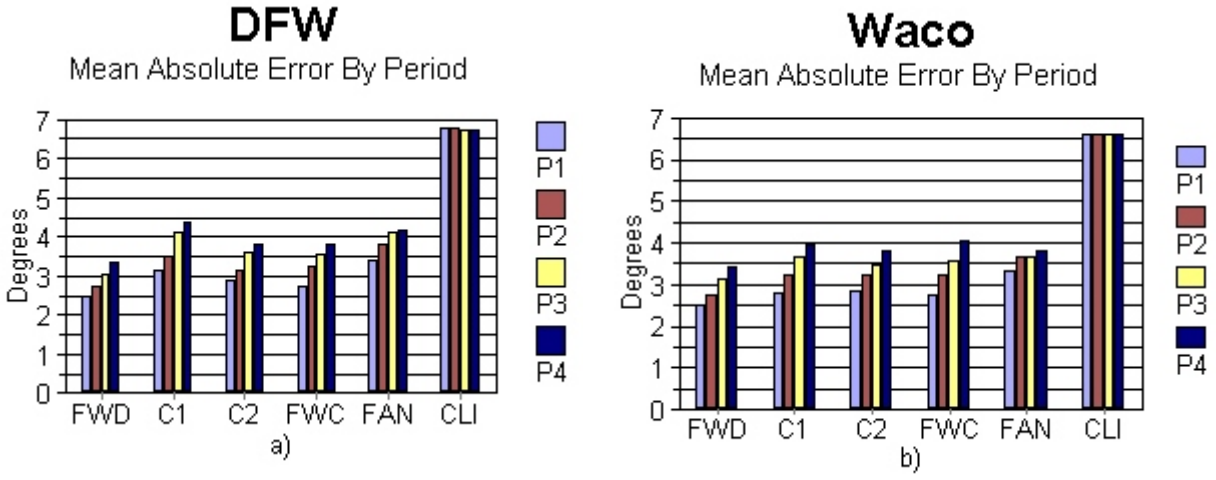


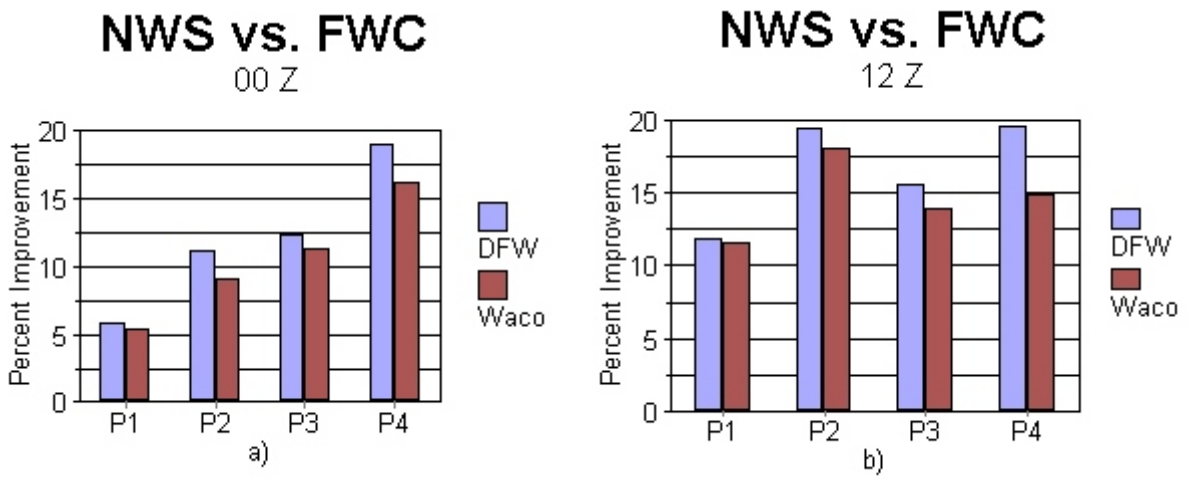
Fig 2. Mean absolute error in degrees Fahrenheit by forecast period for a.)DFW b.)Waco.

It is apparent from the previous statistics that FWD showed improvement over the other participants in total and mean absolute errors. More detail is seen when the percent improvement is calculated for FWD against all participants individually.

The first comparison over the FWC (Figs. 3a and 3b) show that FWD had the greatest improvement in period 4 at both DFW and Waco, and the least percent improvement in period one. It is interesting that percent improvement nearly doubled in period one and two for the forecasts which used the 12Z data. One hypothesis for this improvement is that forecasters had more time for analysis and model comparison when putting together the forecast using the 12Z data, therefore, a more accurate forecast could be made.

When comparing the percent improvement over the FAN (Figs. 4a and 4b) the largest improvement occurred in the first two periods; however, unlike the FWC scores, the FAN improved with time. It is also interesting that the FAN outperformed FWC at Waco in period 4.

Fig 3. FWD percent improvement over FWC at a) 00Z b) 12Z



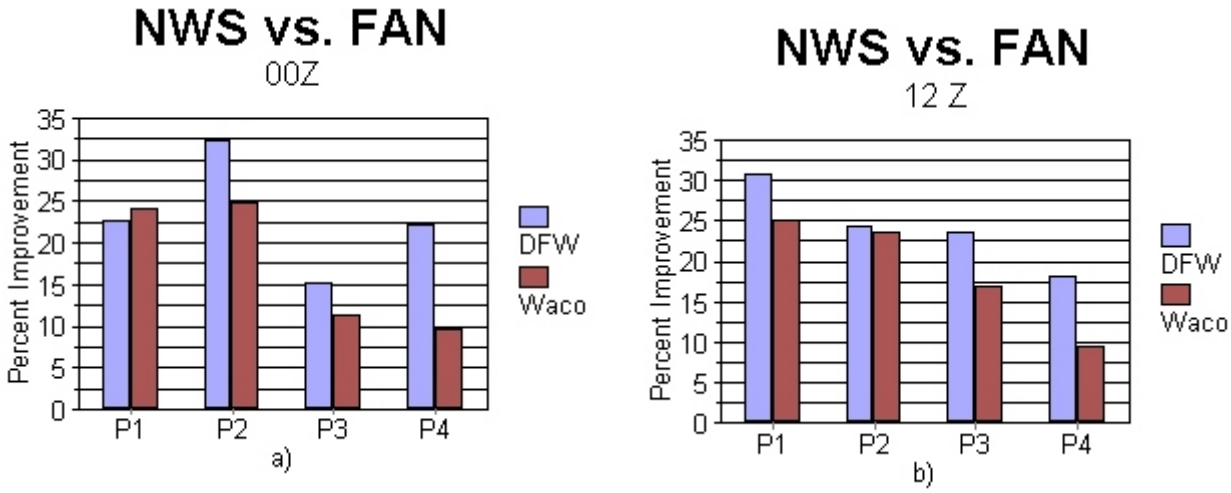


Fig 4. FWD percent improvement over FAN at a) 00Z b) 12Z